

Filtri corretti per serie storiche stagionali con ciclo-trend localmente polinomiale approssimabile

Paolo Chirico (chirico@econ.unito.it)

Dipartimento di Statistica e Matematica applicata alle Scienze Umane,
Università degli Studi di Torino.

Sommario.

Un punto fondamentale nell'analisi dinamica di molti fenomeni economici è l'individuazione della tendenza di fondo del fenomeno stesso. Tale operazione, effettuata sulla base di una serie storica di dati, viene detta filtraggio in quanto consiste nel filtrare il segnale della tendenza da altri effetti di natura congiunturale, stagionale ed accidentale presenti nei dati.

Il presente lavoro tratta appunto il problema del filtraggio del trend nel caso di serie storiche stagionali. L'impostazione, tipicamente temporale, illustra una metodologia per ottenere filtri corretti, che conservano un trend localmente polinomiale-approssimabile (LPA), annullano la componente stagionale e presentano opportune caratteristiche di smussamento (paragrafo 2). La valutazione di tali filtri avviene nella duplice prospettiva temporale e frequenziale, evidenziando punti di collegamento tra i due approcci (paragrafi 3, 4 e 5).

1. Introduzione

L'impostazione classica dell'analisi delle serie storiche si basa sulla scomposizione della dinamica di una serie in due movimenti fondamentali:

- uno rispondente agli effetti di lungo, medio e breve periodo agenti sul fenomeno, detto componente *sistematica*, in quanto si assume che detti effetti possano essere spiegati in funzione del tempo; da qui la denominazione di impostazione temporale;
- uno rispondente agli effetti che agiscono occasionalmente in ogni tempo e pertanto detto componente *accidentale*.

La parte sistematica viene a sua volta distinta in tre movimenti:

1) Uno, detto *trend* f_t , che identifica la tendenza di fondo della serie. Tale movimento generalmente descrive una traiettoria abbastanza smussata (priva di curvature strette) e pertanto è approssimabile, in qualsiasi intervallo temporale locale $[t-m_1, t+m_2]$ da una funzione polinomiale di opportuno grado p :

$$f_{t+h} \approx f_t + b_1(t)h + b_2(t)h^2 + \dots + b_p(t)h^p \quad 2)$$

Il trend risponde agli effetti di lungo periodo e da un punto di vista frequenziale, può essere visto come il miscuglio di armoniche:

$$a_j \cos(\lambda_j t) + b_j \sin(\lambda_j t)$$

principalmente di ciclo lungo, ovvero a bassa frequenza λ_j .

¹ I coefficienti $b(t)$ della funzione polinomiale variano allo scorrere della finestra temporale, altrimenti si dovrebbe assumere un'unica approssimazione polinomiale su tutto l'arco temporale.

2) Un eventuale movimento di medio termine, oscillatorio di periodicità solo in parte regolare, detto *ciclo* c_t , che descrive gli effetti ciclico-congiunturali eventualmente presenti nel fenomeno allo studio; da un punto di vista frequenziale è caratterizzato principalmente da armoniche di banda media. In molti fenomeni di natura economica il ciclo influenza il trend di fondo rendendo impossibile una trattazione distinta dei due movimenti, che per tale ragione vengono fusi in un unico “*ciclo-trend*” o trend di medio periodo m_t . Anche il ciclo-trend può essere localmente polinomiale approssimabile.

3) Se il fenomeno è considerato con cadenza inferiore all'anno, è solitamente presente un movimento stagionale s_t oscillatorio, abbastanza regolare di periodo s (breve periodo), tale che:

$$\sum_{h=0}^{s-1} s_{t+h} \approx 0^2 \quad 3)$$

E' caratterizzato da armoniche con frequenza nell'intorno della frequenza tipica $1/s$.

Il movimento residuale, descrivibile per differenza tra dato storico e i tre movimenti citati, è tipicamente oscillatorio irregolare. Per tale ragione è inteso come il risultato di effetti aleatori occasionali influenti sul fenomeno e sintetizzabili in un processo stocastico a media nulla, detta componente accidentale a_t . Per la sua natura è formalizzabile con un'infinità di armoniche di ampiezza casuale e frequenza su tutta la banda.

L'individuazione delle singole componenti è l'obiettivo finale dell'analisi e solitamente questa avviene partendo dalla determinazione del trend o del ciclo-trend.

Nei paragrafi a seguire verrà dapprima illustrata una metodologia per ottenere filtri corretti, che conservino ciclo-trend localmente polinomiale-approssimabili filtrando serie storiche stagionali; successivamente si procederà ad un'analisi critica e comparativa di tale tecnica.

2. Medie mobili ottimali.

Ipotizzando una scomposizione, con trend di medio periodo LPA:

$$y_t = m_t + s_t + a_t$$

una media mobile (m.m) $\theta' = [\theta_{-m_1}, \dots, \theta_0, \dots, \theta_{m_2}]$ (di ordine $d = m_1 + m_2 + 1$ ³) che filtri correttamente il trend m_t deve essere tale che:

$$\sum_{h=-m_1}^{m_2} \theta_h m_{t+h} = m_t \quad 4)$$

$$\sum_{h=-m_1}^{m_2} \theta_h s_{t+h} = 0 \quad 4bis)$$

ovvero conservi il trend e annulli la stagionalità. Sostituendo la 2) (intesa per m_t) nella 4) con qualche passaggio si ottiene:

$$\sum_{h=-m_1}^{m_2} \theta_h m_{t+h} = m_t \left[\sum_{h=-m_1}^{m_2} \theta_h \right] + b_1(t) \left[\sum_{h=-m_1}^{m_2} h \theta_h \right] + \dots + b_p(t) \left[\sum_{h=-m_1}^{m_2} h^p \theta_h \right] = m_t \quad 5)$$

² Il segno di approssimazione permette di contemplare fluttuazioni stagionali che evolvono nel tempo.

³ m_1 o m_2 possono essere pari a 0, ma non contemporaneamente.

Tale condizione risulta soddisfatta se la prima grandezza tra parentesi quadre è pari a 1 e le altre sono pari a 0. In forma matriciale:

$$\mathbf{T}\boldsymbol{\theta}=\mathbf{e}_1 \quad 6)$$

$$\text{dove } \mathbf{T} = \begin{bmatrix} 1 & 1 & \dots & 1 & \dots & 1 & 1 \\ -m_1 & -m_1+1 & \dots & 0 & \dots & m_2-1 & m_2 \\ \vdots & & \dots & 0 & \dots & & \vdots \\ (-m_1)^p & & \dots & 0 & \dots & & (m_2)^p \end{bmatrix} \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

La condizione 4bis) è soddisfatta se $\boldsymbol{\theta}$ è la composizione di due medie mobili \mathbf{w} e \mathbf{s} , di cui $\mathbf{s}' = [1/s, \dots, 1/s]$ è una media mobile di ordine s che annulla una fluttuazione ciclica di periodo s^4 . Pertanto:

$$\boldsymbol{\theta} = \mathbf{S}\mathbf{w} \quad \text{dove } \mathbf{S} = \begin{bmatrix} 1/s & 0 & \dots & 0 \\ 1/s & 1/s & & \vdots \\ \vdots & \vdots & & 0 \\ 1/s & & & 1/s \\ 0 & 1/s & & \\ \vdots & 0 & & \vdots \\ 0 & 0 & & 1/s \end{bmatrix} \quad \mathbf{w}' = [w_1, w_2, \dots, w_k] \quad 7)$$

Oltre che corretta è auspicabile che la m.m. sia determinata secondo criteri di efficienza e smussamento. Se il trend da filtrare è di lungo periodo.

A tale fine si è adottato il criterio di Henderson, ovvero di minimizzare l'indice di smussamento:

$$S = \sum_{h=-m_1}^{m_2+3} (\Delta^3 \theta_h)^2 \quad 8)$$

S indica la capacità di filtrare un andamento per sua natura smussato, quale un trend; detta capacità è tanto maggiore quanto minore è S .

Pertanto $\boldsymbol{\theta}$ è determinato trovando \mathbf{w} , soluzione della seguente ottimizzazione:

$$\begin{cases} \min_{\mathbf{w}} \mathbf{w}' \mathbf{S}' \mathbf{H}' \mathbf{H} \mathbf{S} \mathbf{w} \\ \mathbf{T} \mathbf{S} \mathbf{w} = \mathbf{e}_1 \end{cases} \quad \text{dove } \mathbf{H} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -3 & 1 & & \vdots \\ 3 & -3 & & 0 \\ -1 & 3 & & 1 \\ 0 & -1 & & -3 \\ \vdots & 0 & & 3 \\ 0 & 0 & & -1 \end{bmatrix} \quad 9)$$

\mathbf{H} è una matrice di dimensione $(d+3 \times d)$. La soluzione è:

$$\mathbf{w} = \mathbf{A}^{-1} \mathbf{B}' (\mathbf{B} \mathbf{A}^{-1} \mathbf{B}')^{-1} \mathbf{e}_1$$

essendo $\mathbf{A} = \mathbf{S}' \mathbf{H}' \mathbf{H} \mathbf{S}$ e $\mathbf{B}' = \mathbf{T} \mathbf{S}$, da cui $\boldsymbol{\theta} = \mathbf{S} \mathbf{w}$

⁴ In realtà qualsiasi m.m. di lunghezza s ad elementi costanti andrebbe bene, anche: $[1, 1, \dots, 1]$.

⁵ \mathbf{S} è una matrice di dimensione $(s+k-1 \times k)$, essendo k la lunghezza di \mathbf{w} .

⁶ Nel computo delle differenze Δ^3 sia $\theta_h=0$ se $h < -m_1$ o $h > m_2$

Generalmente la minimizzazione di 8) comporta anche buoni risultati in termini di efficienza, intesa come riduzione della variabilità degli effetti accidentali. Se la componente accidentale è un White-Noise, ovvero un processo stocastico incorrelato, tale riduzione di variabilità è data da:

$$VR = \sum_{h=-m1}^{m2} \theta_h^2$$

in quanto
$$Var \left[\sum_{h=-m1}^{m2} \theta_h a_{t-h} \right] = \sigma_a^2 \sum_{h=-m1}^{m2} \theta_h^2 \quad \text{essendo } \sigma_a^2 = Var[a_t]$$

Benché l'assunzione di incorrelazione non sia realistica per la maggior parte delle serie socio-economiche, l'approssimazione della riduzione della varianza residua a tale indicatore è comunque accettabile.

3. Calibratura della m.m.

In quanto esposto si evince una caratteristica generale delle m.m.: i loro coefficienti non dipendono dai coefficienti del polinomio che conservano, ma dal suo grado e dalla lunghezza dell'intervallo su cui vale l'approssimazione. Due sono quindi i fattori per determinare la m.m. più opportuna al trend che si desidera filtrare:

- la lunghezza d della m.m.
- il grado p del polinomio da conservare;

Ovviamente questi due fattori vanno scelti in funzione delle caratteristiche del fenomeno allo studio e del contesto temporale d'analisi: se si analizza la tendenza di medio periodo del PIL di un paese con dati trimestrali, la lunghezza della m.m dovrebbe essere da 17 a 19 (quattro-cinque anni); il trend dovrebbe essere abbastanza smussato e quindi la sua approssimazione polinomiale sarà di 2° grado. Per contro il trend di medio periodo nel caso delle vendite di automobili dovrebbe riguardare un ambito temporale più ristretto (11 o 13 trimestri) ed essendo di logica meno smussato, l'approssimazione polinomiale dovrebbe essere di 3°.

In generale un aumento (diminuzione) di grado significa un vincolo in più (in meno) nella determinazione della m.m e quindi:

- una minore (maggiore) capacità di smussamento;
- una minore (maggiore) riduzione della varianza residua.

A livello frequenziale questo significa un filtro meno (più) passa basso, essendo il trend polinomiale da conservare meno (più) caratterizzato da frequenze basse.

Allo stesso modo una maggiore (minore) lunghezza della m.m. significa più (meno) gradi di libertà nella determinazione della m.m e quindi:

- una maggiore (minore) capacità di smussamento;
- una maggiore (minore) riduzione della varianza residua.

A livello frequenziale questo significa un filtro più (meno) passa basso, in quanto l'ampliamento (restrizione) dell'intervallo di approssimazione incide ovviamente sul maggiore (minore) smussamento del trend.

Di norma i valori di d e p utilizzati nelle applicazioni non creano problemi alla determinazione delle m.m., tuttavia va ricordato, che esiste una lunghezza minima delle m.m., che dipende dal grado del polinomio approssimante, dal tipo di stagionalità e dal fatto che la media sia simmetrica o meno. Infatti dalla 7) si deduce che la lunghezza della m.m è $d = s+k-1$, dove k è il numero di coefficienti liberi (gradi di libertà) per determinare la m.m.

Se la m.m. non è centrata simmetrica, i vincoli da soddisfare sono $p+1$, mentre i gradi di libertà sono k , quindi $k \geq p+1$. Se invece la m.m. è simmetrica centrata, allora bisogna tenere conto delle seguenti proposizioni:

- se s è pari, allora k deve essere dispari e viceversa (perché d deve essere sempre dispari);
- se k è dispari i gradi di libertà sono $(k+1)/2$, altrimenti sono $k/2$;
- se p è dispari i vincoli da soddisfare sono $(p+1)/2$, altrimenti sono $1+p/2$;
- i gradi di libertà devono essere almeno pari ai vincoli.

4. Alcune m.m a confronto.

Con la metodologia descritta si sono derivate le seguenti m.m centrate per dati trimestrali e trend di medio periodo LPA

- M9,3 di lunghezza 9, conservante trend polinomiale di 3° grado;
- M9,4 di lunghezza 9, conservante trend polinomiale di 4° grado;
- M11,3 di lunghezza 11, conservante trend polinomiale di 3° grado;

che sono state confrontate con la media mobile di Henderson simmetrica di lunghezza 9 (H9). Tale m.m., come noto, è derivata per conservare in generale trend polinomiali di 2° grado, tuttavia quando è centrata conserva anche trend di 3° grado.

Tab 1: Coefficienti delle m.m considerate.

m.m.	t	t-1	t-2	t-3	t-4	t-5
H9	0,331	0,267	0,118	-0,010	-0,041	
M9,3	0,326	0,268	0,128	-0,016	-0,040	
M9,4	0,141	0,453	0,125	-0,203	0,055	
M11,3	0,301	0,248	0,125	0,024	-0,026	-0,022

Come primo confronto si sono considerati i seguenti indici:

- indice di smussamento
$$S = \sum_{h=-m_1}^{m_2+3} (\Delta^3 \theta_h)^2$$
- riduzione della varianza residua
$$VR = \sum_{h=-m_1}^{m_2} \theta_h^2$$
- indice di distorsione locale
$$l-bias = 1 - \sum_{h=-t_1}^{t_2} \theta_h$$
- indice di distorsione stagionale:
$$s-bias = H(1/4)^8$$
- errore quadratico medio locale
$$LMSE = VR + (l-bias)^2$$

Il significato dei primi due indici è già stato illustrato; il terzo indice è definito da Dagum e Luati⁹ *distorsione locale* e indica la tendenza della m.m. a sovrastimare (valori negativi) o sottostimare (valori positivi) il segnale o trend. L'idea di base, infatti, è che un filtro con coefficienti equidistribuiti abbia la tendenza a smussare eccessivamente il segnale generando una distorsione.

⁷ In quanto $\sum_{h=-m_1}^{m_1} h^r \theta_h = 0$ per r dispari

⁸ $H(1/4)$ è la funzione di guadagno per la frequenza stagionale (si veda oltre).

⁹ Dagum E.B., Luati A. (2004).

Di conseguenza è preferibile che un maggiore peso abbiano i coefficienti centrali, ma non troppo, altrimenti si ricade nel difetto opposto. L'indice informa in tal senso.

Il quarto indicatore dice a quanto viene ridotta la fluttuazione stagionale. Questa, infatti, può essere rappresentata come una somma di armoniche di frequenza pari a $1/4$ (nel caso di dati trimestrali) ed il valore della funzione di guadagno $H(\cdot)$ per tale frequenza indica il coefficiente di riduzione delle corrispondenti armoniche.

Il quinto indicatore, proposto anch'esso da Dagum e Luati, fornisce l'errore quadratico medio delle stime del trend in considerazione della distorsione locale. Per questo è calcolato come riduzione della varianza residua più distorsione locale al quadrato.

Tab 2: Gli indici sulle medie considerate.

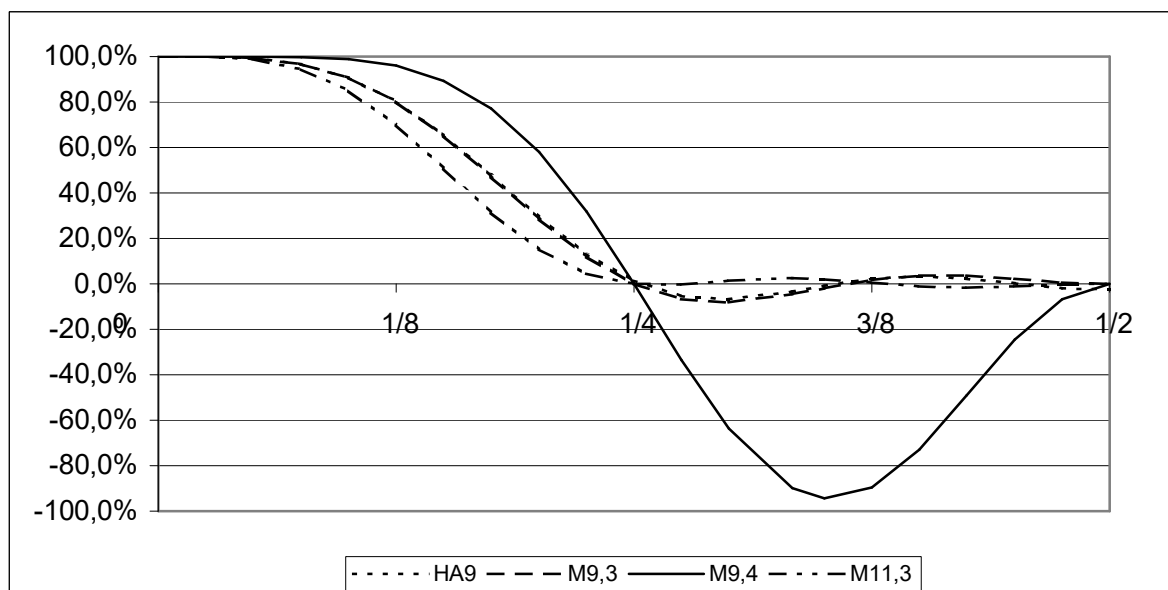
indice	H9	M9,3	M9,4	M11,3
S	0,067	0,073	6,601	0,028
VR	28,3%	28,3%	55,0%	24,8%
lbias	13,6%	14,3%	-4,7%	20,3%
sbias	1,3%	0,0%	0,0%	0,0%
LMSE	30,2%	30,3%	55,2%	29,0%

Sulla base di tali indicatori, risulta una forte similitudine tra M9,3 ed H9 (già evidente a livello di coefficienti); per altro H9, contrariamente a M9,3, presenta una lieve distorsione stagionale (1,3%) e quindi non può considerarsi formalmente corretto. Per quanto riguarda M9,4, la sua capacità di smussamento è molto bassa, come pure la riduzione della varianza: con tali valori non sembra essere interessante a livello pratico. Discorso diverso per M11,3: capacità di smussamento e riduzione della varianza sono abbastanza buoni.

4.1. Confronti frequenziali.

Successivamente si è valutato sulle m.m. il profilo della funzione di guadagno $H(\lambda)$, che come noto indica, per ogni possibile fluttuazione (armonica) di frequenza λ , presente nel processo, l'effetto di riduzione (smussamento) della fluttuazione prodotto dalla m.m.

Graf. 1: Funzioni di guadagno a confronto.



L'analisi della funzione di guadagno è importante perché in base ad essa si può valutare la capacità di una m.m. di filtrare certi trend. Infatti tutte e quattro le m.m. considerate lasciano passare un segnale esattamente polinomiale di 3° grado, ma occorre ricordare che il ciclo-trend è solo in via

approssimata localmente polinomiale. Pertanto è opportuno che ognuna di queste m.m. non sia troppo selettiva rispetto a lievi differenze nella dinamica polinomiale. Tale caratteristica è assicurata se la m.m. lascia maggiormente passare quelle frequenze che maggiormente caratterizzano il ciclo-trend m_t^{10} . La m.m. deve anche annullare l'effetto stagionale e quindi annullare innanzitutto la frequenza stagionale ($\lambda=1/4$), smorzando il più possibile nel suo intorno¹¹.

Anche le frequenze alte devono essere ridotte al massimo, in quanto entrano nella componente accidentale, che ovviamente va smorzata il più possibile.

L'analisi grafica di $H(\lambda)$ ribadisce la stretta somiglianza tra M9,3 ed H9. Conferma inoltre i limiti di M9,4 che lascia passare maggiormente le frequenze medie e le frequenze alte (con l'eccezione, ovviamente di $\lambda=1/4$). Tale circostanza la rende più sensibile quindi ad effetti di breve periodo e di natura accidentale. Teoricamente potrebbe essere indicata per cogliere trend di medio periodo m_t poco lisciati, ma solo se gli effetti accidentali (rumore) siano contenuti rispetto al segnale. Per contro M11,3 presenta caratteristiche decisamente opposte: blocca quasi del tutto le frequenze alte e lascia passare solo frequenze basse. Per tali ragioni sembra indicata per ciclo-trend relativamente smussati.

E' facile notare come tali deduzioni siano del tutto coerenti con i dati riportati in tab 2.

5. Conclusioni.

L'impostazione seguita per derivare le m.m. è tipicamente temporale in quanto si basa:

- sulla presunta approssimazione del trend ad una funzione polinomiale del tempo;
- sull'assunzione che la somma delle fluttuazioni stagionali in un ciclo stagionale sia approssimativamente nulla.

Tale impostazione non richiede che l'approssimazione del trend ad una funzione polinomiale sia su tutto l'arco temporale, ma solo sugli intervalli di definizione delle m.m.. In questo modo il trend non è complessivamente vincolato ad un unico polinomio, bensì è vincolato in termini frequenziali: deve essere (matematicamente) composto da un certo mix di armoniche per garantire l'approssimazione sugli intervalli mobili dei filtri. In termini molto generali la regola può essere sintetizzata in:

- intervallo più (meno) ampio → maggiore (minore) presenza di armoniche di banda bassa;
- grado maggiore (minore) → maggiore (minore) presenza di armoniche di banda medio-alta.

Non dipendendo le m.m. dai coefficienti del polinomio approssimante e potendo questi (moderatamente, ma anche progressivamente) variare allo scorrere dell'intervallo, dette m.m. diventano filtri di segnali vincolati essenzialmente in termini frequenziali.

Considerazioni analoghe valgono per la componente stagionale. Infatti assumere che la somma delle fluttuazioni stagionali sia, ogni s periodi, approssimativamente nulla, significa contemplare anche stagionalità (moderatamente) evolutive e, a livello frequenziale, asserire che le armoniche stagionali si distribuiscono nell'intorno della frequenza $1/s$. Le m.m. proposte, annullando le armoniche di frequenza $1/s$, smorzano generalmente anche le armoniche in tale intorno.

Quanto finora asserito vale ovviamente, se valgono le approssimazioni adottate per quanto riguarda trend e stagionalità. Ovvero se gli errori di approssimazione sono trascurabili e/o ragionevolmente tollerabili. Per molti fenomeni economici tali assunzione hanno senso, ma è bene valutarne preventivamente la fondatezza.

¹⁰ La gamma di frequenze varia dalla tipologia di fenomeno allo studio. Indicativamente può essere (0, 1/10) per molti fenomeni di natura economica.

¹¹ Qualora la componente stagionale sia evolutiva.

In conclusione si ritiene che l'impostazione adottata, pur mantenendo la semplicità formale tipica dell'approccio temporale, fornisca filtri con caratteristiche apprezzabili anche in termini frequenziali.

Bibliografia.

- Chirico P. (1992), Optimal Filtering of the Local Polynomial Trend in Seasonal Time Series , *Atti del Convegno SIS 2002*, Milano.
- Dagum E. B. (1978), Modelling, Forecasting and Seasonal Adjustment Economic Time Series with Arima Method, *Cambridge Conference on Time Series Analysis*.
- Dagum E.B. and Luati A. (2002a): Global and Local Statistical Properties of Fixed-Length Nonparametric Smoothers, *Statistical Methods and Applications*, vol. 11, n. 3.
- Dagum E.B. and Luati A. (2002b): Fitting and Smoothing Properties of Length Constrained Smoothers Applied to Time Series, *Proceedings of Gesellschaft fur Klassifikation*. Germany.
- Dagum E. B. Luati A. (2004), Relation between local and Global Nonparametric estimators measure of Fitting and Smoothing, *Studies in Nonlinear Dynamics & Econometrics*.
- Di Fonzo T., Lisi F. (2005), *Serie Storiche Economiche*, Carocci, Roma.
- Faliva M. (1978), Optimal filtering for seasonal Adjustment of Quarterly Data, *Rivista Internazionale di scienze sociali*, LXXXVI, 1.
- Henderson R. (1916): Note on Graduation by Adjusted Average, *Transaction of the Actuarial Society of America*, 17
- Passamani G. (1986), *Analisi Classica delle componenti di Trend e Periodiche di una Serie Storica*, Mimeo, Padova.
- Zani S. (1982), *Indicatori statistici della congiuntura*, Loescher, Torino